

Lior Pachter

Associate Professor

Departments of Mathematics, Molecular & Cellular Biology
and Computer Science

University of California, Berkeley
Berkeley, CA 94720-3840, USA

Phone: (510) 642-2028

Fax: (510) 642-8204

lpachter@math.berkeley.edu

<http://math.berkeley.edu/~lpachter/>

Research Interests

- **Computational Biology:** Alignment, Phylogeny, Functional and Comparative Genomics
Metagenomics, Population Genetics, Genome Dynamics and Regulation
- **Mathematics:** Combinatorics, Algebra, Computational Geometry
- **Computer Science:** Algorithms, Machine Learning, Scientific Computing
- **Statistics:** Stochastic Processes, Algebraic Statistics

Positions Held

University of California, Berkeley (8/99–present)

- Associate Professor of Mathematics, Molecular & Cellular Biology and Computer Science (7/08–present)
 - Member, Center for Computational Biology
 - Member, Center for Emerging and Neglected Diseases
 - Member, Center for Integrative Genomics
 - Member, Graduate Group in Computational and Genomic Biology
 - Member, Graduate Group in Computational Science and Engineering
 - Member, Joint UCSF/Berkeley Graduate Program in Bioengineering
 - Member, QB3 California Institute for Quantitative Biomedical Research
- Associate Professor of Mathematics & Computer Science (7/06–7/08)
- Associate Professor of Mathematics (7/05–7/06)
- Assistant Professor of Mathematics (8/01–6/05)
- Visiting Assistant Professor of Mathematics (8/99–8/01)

University of Oxford (9/06–07/07)

- Visiting Professor (while on sabbatical)
 - Visitor, Mathematical Institute, hosted by Professor Philip K. Maini
 - Visitor, Department of Statistics, hosted by Professor Jotun Hein

Education

Massachusetts Institute of Technology (9/94–6/99)

- PhD in Mathematics (6/99) Advisor: Professor Bonnie A. Berger
Co-advisors Professor Eric S. Lander and Professor Daniel J. Kleitman.

California Institute of Technology (9/90–6/94)

- B.S. in Mathematics (6/94).
Caltech merit award for academic achievement.

Honors and Fellowships

- Miller Research Professorship (Fall 2009)
- Winner of best paper award for “multiple alignment by sequence annealing”,
(joint with Ariel Schwartz).
2006 European Conference on Computational Biology.
- 2004 National Science Foundation Faculty Early Career Development
(CAREER) Award.
- 2003–2004 Sloan Research Fellow
computational and evolutionary molecular biology.
- 2003 Federal Laboratory Consortium for Technology Transfer Award.
For the VISTA and AVID software: As of 2004, Over 2000 free academic software
licenses distributed to 53 countries, and online web use by 4389 researchers
from 46 countries submitting more than 51,000 sequences for analysis. Eight
commercial licenses sold to pharmaceutical companies..

Grants

- Laboratory for Mathematical and Computational Biology
3/08–6/11 (\$500,000), private donor.
- Fundamental laws of biology, co-PI (with Bernd Sturmfels)
9/05–3/09 (\$750,000), DARPA.
- CAREER: Comparison and Annotation of Multiple Whole Genomes, PI
6/04–6/09 (\$400,000), NSF.
- GENCODE: The Encyclopedia of Genes and Gene Variants (main PI: Roderic Guigó)
10/03–10/06 (\$72,000), NIH.
- Cross-species gene finding and annotation, PI
6/02–6/05 (\$928,728), NIH (R01).

Publications

Books

1. L. Pachter and B. Sturmfels: Algebraic Statistics for Computational Biology, Cambridge University Press, October 2005.

Submitted Manuscripts

82. A. Lapuk et al. : Micorarrays and next generation sequencing identifies alternatively spliced markers and therapeutic targets in breast cancer, submitted.
81. M. Hilty, C. Burke, H. Pedro, A. Ervine, L. Poulter, L. Pachter, M.F. Moffatt and W.O.C. Cookson: Molecular detection of differences in airway microbial communities in normal adults and subjects with asthma and chronic obstructive pulmonary disease, submitted.
80. R. Bradley, A. Roberts, M. Smoot, S. Juvekar, J. Do, C. Dewey, I. Holmes and L. Pachter: Fast statistical alignment, submitted.
79. A. Caspi and L. Pachter: Ancestral inference with phylogenetic profiles in the presence of observation error, submitted.
78. P. Huggins and L. Pachter: Selecting universities: personal preference and rankings, submitted.
77. C. Dewey, A. Schwartz, N. Bray and L. Pachter: Finding the trees in Darwin's forest, submitted.

Journal Publications

76. J. Morton, L. Pachter, A. Shiu, B. Sturmfels and O. Wienand: Convex rank tests and semigraphoids, *SIAM Journal on Discrete Mathematics*, in press.
75. D. Levy and L. Pachter: The neighbor-net algorithm, *Advances in Applied Mathematics*, in press.
76. R. Bradley, L. Pachter and I. Holmes: Specific alignment of structured RNA: stochastic grammars and sequence annealing, *Bioinformatics*, 24 (2008) 2677–2683.
75. R. Mihaescu and L. Pachter: Combinatorics of least squares trees, *Proceedings of the National Academy of Sciences*, 105 (2008) p 13206–13211.
74. M-L. Dequéant et al.: Comparing pattern detection methods in microarray time series of the segmentation clock, *PLoS One*, 3 (2008) e2856.

73. K. Eickmeyer, P. Huggins, L. Pachter and R. Yoshida: On the optimality of neighbor-joining, *Algorithms for Molecular Biology*, 3 (2008).
72. N. Eriksson, L. Pachter, S.Y. Rhee, Y. Mitsuya, R.W. Shafer and N. Beerenwinkel: Viral population estimation using pyrosequencing, *PLoS Computational Biology*, Volume 4 (2008) e1000074.
71. R. Satija, L. Pachter and J. Hein: Combining statistical alignment and phylogenetic footprinting to detect regulatory elements, *Bioinformatics*, advanced access (2008).
70. D. Levy, R. Mihaescu and L. Pachter: Why neighbor joining works, *Algorithmica*, online first (2008).
69. A. Stark et al.: Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures, *Nature*, Volume 450 (2007) p 219–232.
68. Drosophila Comparative Genome Sequencing and Analysis Consortium: Evolution of genes and genomes in the context of the Drosophila phylogeny, *Nature*, Volume 45 (2007) p 203–218.
67. D.J. Begun et al.: Population genomics: whole-genome analysis of polymorphism and divergence in Drosophila simulans, *PLoS Biology*, Volume 5 Number 11 (2007) e310.
66. P. Huggins, L. Pachter and B. Sturmfels: Towards the human genotype, *Bulletin of Mathematical Biology*, Volume 69 Number 8 (2007) p 2723–2725.
65. N. Beerenwinkel, L. Pachter and B. Sturmfels: Epistasis and the shapes of fitness landscapes, *Statistica Sinica*, Volume 17 Number 4 (2007) p 1317–1342.
64. J. Morton, L. Pachter, A. Shiu and B. Sturmfels: The cyclohedron test for finding periodic genes in time course expression studies, *Statistical Applications in Genetics and Molecular Biology*, Volume 6 Issue 1 (2007).
63. The ENCODE Project Consortium: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project, *Nature*, 447 (2007) p 799–816.
62. E.H. Margulies *et al.*: Analysis of deep mammalian sequence alignments and constraint predictions for 1% of the human genome, *Genome Research*, 17 (2007) p 760–774.
61. S. Chatterji and L. Pachter: Patterns of gene duplication in the ENCODE regions suggest a confounding factor, *Genomics*, 90 (2007) p 44–48.
60. L. Pachter: Interpreting the unculturable majority, *Nature Methods*, 4 (2007) p 479–480.

59. N. Beerenwinkel, L. Pachter, B. Sturmfels, S. Elena and R. Lenski: Analysis of epistatic interactions and fitness landscapes using a new geometric approach, *BMC Evolutionary Biology*, 7:60 (2007).
58. L. Pachter and B. Sturmfels: The mathematics of phylogenomics, *SIAM review*, 49 (2007) p 3–31.
57. C. Dewey, P. Huggins, K. Woods, B. Sturmfels and L. Pachter, Parametric alignment of Drosophila genomes: *PLoS Computational Biology*, 2 (2006) e73.
56. C. Dewey and L. Pachter: Evolution at the Nucleotide Level: The Problem of Multiple Whole Genome Alignment, *Human Molecular Genetics*, 15 (2006) R51 – R56.
55. S. Chatterji and L. Pachter, Reference based annotation with Gene-Mapper: *Genome Biology*, 7 (2006) R29.
54. D. Levy, L. Pachter and R. Yoshida: Beyond pairwise distances: neighbor joining with phylogenetic diversity estimates, *Molecular Biology and Evolution*, 23 (2006) p 491 – 498.
53. S. Lall, D. Grün, A. Krek, K. Chen, Y. Wang, C. Dewey, P. Sood, T. Colombo, N. Bray, P. MacMenamin, H. Kao, K.C. Gunsalus, L. Pachter, F. Piano and N. Rajewsky: A genome-wide map of conserved microRNA targets in *C. elegans*, *Current Biology*, 16 (2006) p 460 – 471.
52. A. Caspi and L. Pachter: Identification of transposable elements using multiple alignments of related genomes, *Genome Research*, 16 (2006) p 260 – 270.
51. K. Chen and L. Pachter: Bioinformatics for whole-genome shotgun sequencing of microbial communities, *PloS Computational Biology*, 1 (2005) p e24.
50. S. Chatterji and L. Pachter: Multiple organism gene finding by collapsed Gibbs sampling, *Journal of Computational Biology*, 12 (2005) p 599 – 608.
49. J. D. McAuliffe, M. I. Jordan and L. Pachter: Subtree power analysis finds optimal species for comparative genomics, *Proceedings of the National Academy of Sciences*, 102 (2005) p 7900 – 7905.
48. International Chicken Genome Sequencing Consortium: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution, *Nature*, 432 (2004) p 695 – 716.
47. D. Boffelli, C. V. Weer, L. Weng, K.D. Lewis, M. I. Shoukry, L. Pachter, D. N. Keys and E. M. Rubin: Intra-species sequence comparisons for annotating genomes, *Genome Research*, 14 (2004) p 695 – 716.

46. K. A. Frazer, L. Pachter, A. V. Poliakov, E. M. Rubin and I. Dubchak: VISTA–computational tools for comparative genomics, *Nucleic Acids Research*, 32 (2004) p 273 – 279.
45. L. Pachter and B. Sturmfels: Parametric Inference for Biological Sequence Analysis, *Proceedings of the National Academy of Sciences*, 101 (2004) p 16138 – 16143.
44. L. Pachter and B. Sturmfels: Tropical Geometry of Statistical Models, *Proceedings of the National Academy of Sciences*, 101 (2004) p 16132 – 16137.
43. The ENCODE Project Consortium: The ENCODE (ENCyclopedia of DNA Elements) project, *Science*, 306 (2004) p 636 – 640.
42. L. Pachter and D. Speyer: Reconstructing trees from subtree weights, *Applied Mathematics Letters*, 17 (2004) p 615 – 621.
41. J. D. McAuliffe, L. Pachter and M. I. Jordan: Multiple-sequence functional annotation and the generalized hidden Markov phylogeny, *Bioinformatics*, 20 (2004) p 1850 – 1860.
40. Rat Genome Sequencing Consortium: Evolution of the Mammalian Genome: Sequence of the Genome of the Brown Norway Rat, *Nature*, vol. 428 (2004), p 493–521 .
39. N. Bray and L. Pachter: MAVID: Constrained ancestral alignment of multiple sequences, *Genome Research*, 14 (2004), p 693–699.
38. K. Chakrabarti and L. Pachter: Visualization of multiple genome annotations and alignments with the K–BROWSER , *Genome Research*, 14 (2004), p 716–720.
37. C. Dewey, J.Q. Wu, S. Cawley, M. Alexandersson, R. Gibbs and L. Pachter: Accurate identification of novel human genes through simultaneous gene prediction in human, mouse, and rat, *Genome Research*, 14 (2004), p 661–664.
36. V.B. Yap and L. Pachter: Identification of evolutionary hotspots in the rodent genomes, *Genome Research*, 14 (2004), p 574–579.
35. S. Cawley and L. Pachter: HMM sampling and applications to gene finding and alternative splicing, *Bioinformatics*, 19 Supplement 2 (2003) p 36–41.
34. F. Lam and L. Pachter: Forcing Numbers for Stop Signs, *Theoretical Computer Science*, 303 (2003) p 409 – 416.
33. N. Bray and L. Pachter: MAVID multiple alignment server, *Nucleic Acids Research*, 31 (2003) p 3525–3526.
32. S. Cawley, L. Pachter and M. Alexandersson: SLAM web server for comparative gene finding and alignment, *Nucleic Acids Research*, 13 (2003) p 3507–3509.

31. F. Lam, M. Alexandersson, L. Pachter: Picking Alignments from (Steiner) Trees, *Journal of Computational Biology*, 10 (2003) p 509–520.
30. M. Alexandersson, S. Cawley, L. Pachter: SLAM- Cross-species gene finding and alignment with a generalized pair hidden Markov model, *Genome Research*, 13 (2003) p 496–502.
29. D. Boffelli, J. McAuliffe, D. Ovcharenko, K. D. Lewis, I. Ovcharenko, L. Pachter, E.M. Rubin: Phylogenetic Shadowing of Primate Sequences to Find Functional Regions of the Human Genome, *Science*, Volume 299, Number 5611 (2003), p 1391–1394.
28. N. Bray, I. Dubchak and L. Pachter: AVID: A Global Alignment Program, *Genome Research*, 13 (2003) p 97 – 102.
27. O. Couronne, A. Poliakov, N. Bray, I. Ovcharenko, T. Ishkhanov, D.V. Ryaboy, E. Rubin, L. Pachter and I. Dubchak: Strategies and Tools for Whole Genome Alignments, *Genome Research*, 13 (2003) p 73 – 80.
26. Mouse Genome Sequencing Consortium: Initial sequencing and comparative analysis of the mouse genome, *Nature*, vol. 420 (2002), p 520–562.
25. L. Pachter , M. Alexandersson and S. Cawley: Applications of Generalized Pair Hidden Markov Models to Alignment and Gene Finding Problems, *Journal of Computational Biology*, Volume 9, Number 2, 2002, p 389 – 400.
24. G. Loots, I. Ovcharenko, L. Pachter and I. Dubchak and E. M. Rubin: Comparative sequence-based approach to high-throughput discovery of functional regulatory elements, *Genome Research*, 12 (2002), p 832–839.
23. I. Dubchak and L. Pachter: The computational challenges of applying comparative-based computational methods to whole genomes, *Briefings in Bioinformatics*, March 2002, vol. 3, no. 1, p 18–22.
22. A. Peter *et al.*: Mapping and identification of essential gene functions on the X chromosome of *Drosophila*, *EMBO reports*, vol. 3, no. 1 (2001), p 34–38.
21. P. V. Benos *et al.*: From First Base: The sequence of the tip of the X chromosome of *D. melanogaster*, *Genome Research*, 11 (2001), p 710–730.
20. C. Mayor, M. Brudno, J.R. Schwartz, A. Poliakov, E. M. Rubin, K. A. Frazer, L. Pachter and I. Dubchak: VISTA: Visualizing global DNA sequence alignments of arbitrary length, *Bioinformatics*, 16 (2000), p 1046–1047.
19. I. Dubchak, C. Mayor, M. Brudno, L. Pachter, E. M. Rubin and K. A. Frazer: Active conservation of noncoding sequences revealed by 3-way species comparisons, *Genome Research*, 10 (2000), p 1304–1306.

18. S. Batzoglou, L. Pachter, J. Mesirov, B. Berger and E. S. Lander: Human and Mouse Gene Structure: Comparative Analysis and Application to Exon Prediction, *Genome Research*, 10 (2000), p 950–958.
17. L. Pachter, S. Batzoglou, V. I. Spitkovsky, E. Banks, E. S. Lander, B. Berger and D. J. Kleitman: A dictionary based approach for gene annotation, *Journal of Computational Biology* (1999), Fall-Winter (6)3-4:419–430.
16. L. Pachter and P. Kim: Forcing Matchings on Square Grids, *Discrete Mathematics*, 190 (1998) p 287–294.
15. D. J. Kleitman and L. Pachter: Finding Convex Sets Among Points in the Plane, *Discrete and Computational Geometry*, 19 (3) (1998) p 405–410.
14. L. Pachter: Constructing Status Injective Graphs, *Discrete Applied Mathematics*, 80 (1) (1997), p 107–113.
13. L. Pachter: Combinatorial Approaches and Conjectures for 2-Divisibility Problems Concerning Domino Tilings of Polyominoes, *Electronic Journal of Combinatorics*, 4(1), R29 (1997), 10pp.

Refereed Conference Publications

12. L. Pachter: An introduction to reconstructing ancestral genomes, *Proceedings of Symposia in Applied Mathematics*, AMS Short Course Subseries, Volume 64 (2007), p 1–20.
11. A. Schwartz and L. Pachter: Multiple alignment by sequence annealing, *Proceedings of the European Conference on Computational Biology, Bioinformatics 23* (2007), e24–e29.
10. J. Morton, L. Pachter, A. Shiu, B. Sturmfels and O. Wienand: Geometry of Rank Tests, *Proceedings of the Third European Workshop on Probabilistic Graphical Models*, September 2006.
9. S. Snir and L. Pachter: Phylogenetic Profiling of Insertions and Deletions in Vertebrate Genomes, *Proceedings of the Tenth Annual International Conference on Computational Molecular Biology (RECOMB 2006)*, p 265 – 280.
8. S. Chatterji and L. Pachter: Multiple Organism Gene Finding by Collapsed Gibbs Sampling, *Proceedings of the Eighth Annual International Conference on Computational Molecular Biology (RECOMB 2004)*, p 187–193.
7. L. Pachter, F. Lam, M. Alexandersson: Picking Alignments from (Steiner) Trees, *Proceedings of the Sixth Annual International Conference on Computational Molecular Biology (RECOMB 2002)*, p 246–253.

6. L. Pachter and M. Pachter: Avoiding points in the plane, *Proceedings of the 40th IEEE Conference on Decision and Control* (2001).
5. L. Pachter, M. Alexandersson, S. Cawley: Applications of Generalized Pair Hidden Markov Models to Alignment and Gene Finding Problems, *Proceedings of the Fifth Annual International Conference on Computational Molecular Biology (RECOMB 2001)*, p 241–248.
4. S. Batzoglou, L. Pachter, J. Mesirov, B. Berger, E. S. Lander: Comparative Analysis of Mouse and Human DNA and Applications to Exon Prediction, *Proceedings of the Fourth Annual International Conference on Computational Molecular Biology (RECOMB 2000)*, p 46–53.
3. L. Pachter, S. Batzoglou, V. I. Spitkovsky, W. Beebee, E. S. Lander, B. Berger and D. J. Kleitman: A dictionary based approach for gene annotation, *Proceedings of the Third Annual International Conference on Computational Molecular Biology (RECOMB 1999)*, p 285–294.
2. S. Batzoglou, B. Berger, D. J. Kleitman, Eric S. Lander and L. Pachter: Recent Developments in Computational Gene Recognition, *Documenta Mathematica, Extra Volume ICM 1998 I* (1998) 1 p 649–658.
1. L. Pachter, H. Snevily and B. Voxman: On Pebbling Graphs, *Proceedings of the Twenty-sixth Southeastern International Conference on Combinatorics, Graph Theory and Computing. Congressus Numerantium*, 107 (1995) p 65–80.

Book Chapters

1. M. Alexandersson, N. Bray and L. Pachter: Pair hidden Markov models, special review for the *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics* (L. B. Jorde, P. Little, M. Dunn and S. Subramanian, editors), October 2005.

Other Manuscripts

5. L. Pachter: A Nobel Prize in *Mathematical Biology*? Mathematics Department Newsletter, UC Berkeley, 2006.
4. R. Guigó, E. Birney, M. Brent, E. Dermitzakis, L. Pachter, H.R. Crollius, V. Solovyev, M.Q. Zhang: Needed for completion of the human genome: hypothesis driven experiments and biologically realistic mathematical models, 2004.
3. M. Alexandersson, S. Cawley and L. Pachter: HMMs for Gene Recognition, 2000.
2. N. Kitchloo and L. Pachter: An interesting result about subset sums, 1993.
1. R. Dawson, B. McDonald, J. Mycielski and L. Pachter: Light Traps, 1992.

Professional Activities

- Member, Scientific Advisory Board, Mathematical Biosciences Institute, 11/2008–12/2011.
- MSRI workshop organizer (with Serkan Hosten and Bernd Sturmfels)
on “Algebraic Statistics”, 2008.
- IMA workshop organizer (with Serkan Hosten and Seth Sullivant)
on “Applications of Algebraic Geometry to Biology, Dynamics and Statistics”, 2007.
- MSRI summer graduate workshop organizer (with Reinhard Laubenbacher)
on “Mathematical aspects of computational biology”, 2006.
- Conference organizer (with Bernd Sturmfels and Seth Sullivant)
on “Computational Biology and Algebraic Statistics”, Clay Institute, 2005.
- Session organizer (with Peer Bork) on “Computational Genomics”
Cold Spring Harbor Laboratories meeting on “Biology of Genomes”, 2004.
- Session organizer (with Allan Bradley) on “Comparative Genomics”
Cold Spring Harbor Laboratories meeting on “Genome Sequencing and Biology”, 2002.
- Session organizer (with Sean Eddy) on “Comparative Genomics”
CSHL/Wellcome Trust joint meeting, 2002.
- Session organizer (with Serafim Batzoglou) on “Gene Regulation”
Pacific Symposium on Biocomputing (PSB), 2003.
- Session organizer (with Inna Dubchak) on “Genome-wide Analysis and
Comparative Genomics”, Pacific Symposium on Biocomputing (PSB), 2002.
- Developer of Berkeley graduate course Math 217
Discrete Mathematics for the Life Sciences, 2007–
- Developer of Berkeley undergraduate course Math 127
Mathematical and Computational Methods in Molecular Biology, 1999–

Outreach

- VIGRE lecture, Department of Statistics, UC Berkeley (11/07)
- Palo Alto High School Mathematics Evening (5/05)
- Los Medanos College Computational Biology Lecture (2/05)

Reviewing

- Associate editor for *Bulletin of Mathematical Biology*, *Mathematical Medicine & Biology*.
- Referee for *Science*, *Advances in Applied Mathematics*, *Bioinformatics*, *Discrete Mathematics*, *Electronic Journal of Combinatorics*, *Genome Research*, *Journal of Computational Biology*, *Journal of Theoretical Biology*, *Nature Genetics*, *Nature Methods*, *Bulletin of Mathematical Biology*, *PLoS Genetics*, *Nucleic Acids Research*, *Journal of Symbolic Computation*, *Journal of Molecular Biology*, *Genome Biology*, *Statistica Sinica*, *Journal of Theoretical Biology*, *BMC Bioinformatics*, *Statistical Applications in Genetics and Molecular Biology*, *PLoS Computational Biology*.
- External reviewer for investigator proposals at the Wellcome Trust Sanger Institute.
- Standing Panel Member, 2004–2006.
NIH Biological Data Management and Analysis grant review group.
- Grant reviewer for the NSF.

Committees

- Faculty Appointments Committee, Department of Mathematics, UC Berkeley, 2007–2008.
- Graduate Admissions Committee, Department of Computer Science, UC Berkeley, 2007–2008.
- Committee on the Profession of the American Mathematical Society, 2005–2008.
- Computational Biology Initiative Search Committee, UC Berkeley, 2004–2006.
- Executive committee
Graduate group in computational and molecular biology, UC Berkeley, 2001–present.
- Conference program committees
11th International Conference on Research in Computational Molecular Biology (2007).
First Argentine Conference on Mathematics and Biology (2005).
First IEEE Computer Society Bioinformatics Conference (2002).
Second Workshop on Algorithms in Bioinformatics, (2002).

Invited Talks

Biology/Genetics Departments and Institutes

- University of California at Berkeley (5/08)
What is an alignment?
- University of Pennsylvania (3/08)
Functional Genomics and the Forest of Life
- The National Institutes of Health (5/07)
From Statistical Sampling to Collection and Interpretation of Metagenomics Data
- The Wellcome Trust Sanger Institute (11/06)
The Majority of Divergence between DNA Sequences is due to Indels
- Stowers Institute for Medical Research (3/06)
Geometry of Rank Tests and Applications to Microarray Time Series Analysis
- The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia (7/02)
Whole Genome Alignment and Gene Finding
- Whitehead Institute for Biomedical Research (5/02)
Whole Genome Alignment
- Johns Hopkins University (5/02)
Comparing the human and mouse genomes

Computational Biology Seminars

- Université de Montréal (11/08)
Robert Cedergren Bioinformatics Colloquium
- Bioinformatics Research Seminar, University of California at Santa Cruz (5/08)
What is the neighbor-joining algorithm?
- Forefront of Genomics Colloquium, Genome Center, UC Davis (11/07)
High-throughput population genomics
- Center for Genomic Regulation, Barcelona (5/07)
From Drosophila and Transposable Elements to the Neighbor-Net Algorithm and Phylogenetic Networks
- University of East Anglia (4/07)
From Drosophila and Transposable Elements to the Neighbor-Net Algorithm and Phylogenetic Networks
- North Carolina State University (3/07)
From Drosophila and Transposable Elements to Phylogenetic Networks
- University of Oxford (2/07)
Comparative Genomics of Drosophila
- Stockholm Bioinformatics Center (1/07)
Why Neighbor-Joining Works
- California Institute for Quantitative Biomedical Research Seminar (11/06)
The Majority of Divergence between DNA Sequences is due to Indels
- Center for Genome Sciences, Cuernavaca Mexico (5/06)

- Annotation and Alignment of the Drosophila Genomes
- Ohio State University (4/06)
 - Annotation and Alignment of the Drosophila Genomes
- University of Pennsylvania – Penn Bioinformatics Forum (2/06)
 - Annotation and Alignment of the Drosophila Genomes
- Virginia Bioinformatics Institute (2/05)
 - Gene Finding
- Brown University (11/04)
 - Parametric Sequence Alignment
- University of California, San Francisco – Cancer Center (11/03)
 - Phylogenetic Shadowing
- Boston University (10/03)
 - Phylogenetic Shadowing
- Jackson Laboratories (8/02)
 - Finding genes in human using the mouse and finding genes in mouse using the human

Computer Science Departments

- University of Pennsylvania (4/07)
 - From Drosophila and Transposable Elements to the Neighbor-Net Algorithm and Phylogenetic Networks
- University of California, Berkeley (10/05)
 - Robust Alignment for Drosophila Genomics.
- Iowa State University (8/04)
 - Whole Genome Alignment and Gene Finding
- University of California, Santa Cruz (1/04)
 - Parametric Inference for Biological Sequence Analysis
- University of California, Santa Cruz (1/02)
 - Hidden Markov Models, Alignment, Gene Recognition and Mice
- Johns Hopkins University (10/98)
 - Onion Peeling and Protein Folding

Mathematics Departments

- Fuzzy Vance Lecture at Oberlin College (2/09)
 - Finding the Trees in Darwin's Forest
- Freie Universität, Berlin (4/08)
 - Mathematics of Distance-Based Phylogenetics
- University of Pennsylvania (3/08)
 - Foundations of Distance-Based Phylogenetics
- University of California at Davis (11/07)
 - From Biology to Mathematics: The Neighbor-Joining Algorithm
- University of Minnesota (2/07)
 - From Drosophila and Transposable Elements to Phylogenetic Networks and Associahedra

- University of Utah (3/06)
Annotation and Alignment of the Drosophila Genomes
- Stanford University (5/04)
Parametric Inference for Biological Sequence Analysis
- University of California, Davis (2/01)
Hidden Markov Models, Alignment, Gene Recognition and Mice
- University of California, Berkeley (11/00)
Hidden Markov Models, Alignment, Gene Recognition and Mice
- University of Washington (5/00)
Counting Walks on Tilings
- University of California, San Diego (1/00)
The Erdős-Szekeres Theorem
- University of California, Berkeley (10/99)
The Erdős-Szekeres Theorem
- University of Southern California (10/99)
Comparative Genomics and Gene Recognition
- Iowa State University (3/99)
Dictionary Approaches for Gene Annotation
- Carnegie Mellon University (2/99)
Dictionary Approaches for Gene Annotation
- Johns Hopkins University (10/98)
The Erdős-Szekeres Theorem
- Oberlin College (2/98)
The Erdős-Szekeres Theorem
- University of Idaho (2/98)
The Erdős-Szekeres Theorem
- California Institute of Technology (2/98)
Combinatorial and Computational Approaches to Gene Recognition
- University of Washington (3/94)
Light Traps

Statistics and Biostatistics Departments

- Statistics Seminar, University of Chicago (11/08)
Statistical foundations of distance-based phylogenetics
- University of California Berkeley – Neyman Seminar (11/08)
What is algebraic statistics and what is it good for?
- Biostatistics and Medical Informatics Seminar, University of Wisconsin (4/08)
Statistical foundations of distance-based phylogenetics
- University of Iceland (12/07)
Least squares tree metrics
- University of Washington (11/07)
Statistical foundations of distance-based phylogenetics methods

- University of Oxford (6/07)
What is Algebraic Statistics?
- University of Oxford (10/06)
Towards the Human Genome
- University of Oxford (11/05)
Alignment and Annotation of the Drosophila Genomes
- University of California Berkeley – Neyman Seminar (11/04)
Gene Finding
- Harvard University (1/02)
Hidden Markov Models, Alignment, Gene Recognition and Mice
- University of California Berkeley – Neyman Seminar (3/00)
Onion Peeling and Protein Folding

Conferences and Workshops

- 7th Australia–New Zealand Mathematics Convention (12/08)
Finding the trees in Darwin's forest
- SAMSI Opening Workshop on Algebraic Methods in Systems Biology and Statistics (9/08)
The algebra and statistics of biological sequence alignment
- High Dimensional Statistics in Biology, Newton Institute, Cambridge (4/08)
Functional genomics and the forest of life
- Banbury Meeting, Cold Spring Harbor (3/08)
Tutorial on Algebraic Statistics for Computational Biology
- Information Theory and Applications Workshop (12/07)
The neighbor-net algorithm and the traveling salesman problem
- Bay Area Mathematical Biology Day (11/06)
Towards the Human Genome
- Summer school, Sophus Lie Conference Center, Nordfjordeid (6/06)
Algebraic statistics, tropical geometry and computational biology (6 lectures)
- Gordon Conference on Molecular Evolution (2/06)
Annotation and Alignment of the Drosophila Genomes
- AMS Special Session on Mathematical Biology, San Antonio (1/06)
Reconstructing Ancestral Genomes
- AMS Special Session on Algebraic Statistics, San Antonio (1/06)
The Posterior Map
- First Argentine Conference on Mathematics and Biology (12/05)
An Introduction to Algebraic Statistics for Computational Biology (3 lectures)
- Clay Workshop on Algebraic Statistics and Computational Biology (11/05)
Algebraic Statistics for Computational Biology
- Advanced Bioinformatics Course, Cold Spring Harbor (11/05)
Whole Genome Alignment
- Banff Research Station (8/04)
Parametric Sequence Alignment

- The Biology of the Genomes Meeting, Cold Spring Harbor (5/04)
Phylogenetic Methods in Genomics and Applications to Functional Element Identification
- Snowbird Learning Theory Workshop (4/04)
Parametric Inference for Biological Sequence Analysis
- Banbury Meeting, Cold Spring Harbor (3/04)
Phylogenetic Methods in Genomics and Applications to Functional Element Identification
- Fundacion La Caixa, Barcelona (11/03)
Phylogenetic Shadowing
- Bertinoro Workshop on Computational Biology (6/03)
Mathematics of Phylogenetic Shadowing
- National Institute of Environmental and Health Safety, Toxigenomics meeting (5/03)
Comparative genomics by phylogenetic shadowing
- Frontiers of Grid Computing, Baton Rouge Louisiana (3/03)
Comparing the human and mouse genomes
- Mathematics Workshop, Taranaki New Zealand (1/03)
Mathematics of Gene Finding and Alignment
- Bay Area Discrete Math Day (10/02)
Picking alignments from (Steiner) Trees
- SIAM Conference on Discrete Mathematics (8/02)
Comparing the Human and Mouse Genomes
- Jackson Laboratories Workshop on Comparative Genomics (8/02)
Whole Genome Alignment and Gene Finding
- Human Genetics Meeting, Cold Spring Harbor (5/02)
Comparative Assembly
- Gordon Conference on Human Genetics (8/01)
Hidden Markov Models, Alignment, Gene Recognition and Mice
- PMMB workshop, Tucson (6/01)
Whole Genome Alignment and Gene Finding
- RECOMB satellite meeting on assembly, Los Angeles (5/01)
Comparative Assembly
- PMMB Workshop, Mathematical Sciences Research Institute (6/00).
Combinatorial and Computational Approaches to Gene Recognition

Teaching

Classes

- Genome Project Laboratory (UCB MCB 247, 1/09)
A hands-on graduate course on whole-genome sequencing, assembly and analysis co-taught with Rachel Brem and Michael Eisen.
- Discrete mathematics for the life sciences (UCB MATH 239)
An introductory graduate course at UC Berkeley on discrete mathematics and applications to current topics in biology developed by L. Pachter.
- Mathematical and computational methods in molecular biology (UCB MATH 127, 9/00, 9/01, 9/02, 1/04, 9/07). This is an undergraduate course at UC Berkeley for mathematicians interested in the applications of mathematics to molecular biology developed by L. Pachter.
- Algebraic combinatorics (UCB MATH 249, 9/05, 1/09)
Taught from the notes *Combinatorial Theory*, by Gian-Carlo Rota.
- Graduate Research Seminar in Computational Biology (with Bernd Sturmfels, UCB MATH 290, 9/03, 1/05, 1/06).
- Graduate Research Seminar in Computational Biology (with Niko Beerenwinkel, UCB MATH 290, 9/05).
- Graduate Research Seminar in Computational Biology (with Richard Karp, UCB MATH 290, 1/00, 9/02).
- Topics in applied mathematics: algebraic statistics for computational biology (with Bernd Sturmfels UCB MATH 275, 9/04).
- Enumerative combinatorics (UCB MATH 199, 1/00).
An independent study course with two students. The study was based on Richard Stanley's Enumerative Combinatorics Vol. 1.
- Numerical analysis (UCB MATH 128a, 1/00, 1/01, 1/03).
- Linear algebra (UCB MATH 110, 9/99, 1/02).
- Multivariable calculus (UCB MATH 53, 1/04).
- Freshman calculus (MIT 18.01, 9/95, 9/96, UCB MATH 16B, 1/06).
Designed course and taught approximately 30 students as part of the MIT concourse program.

Short Courses

- Applications of algebraic statistics (5/07)
Course taught at the University of Barcelona, Spain.
- Algebraic statistics for computational biology (3/05)
Course taught at the National University of Cordoba, Argentina.
- Hidden Markov models for alignment, gene finding and protein motif detection
Harvard short course on "Statistics and Genomics", Boston (1/02)
- Multiple Sequence Alignment (11/06)
Lecture series at the Instituto Gulbenkian de Ciencia, Oerias, Portugal.

Research Supervision

Current graduate students

- Nicolas Bray (mathematics).
Ph.D. thesis June 2009.
- Hélder António Martins Pedro (computational biology at the Instituto Gulbenkian)
Ph.D. thesis June 2010.
- Rahul Satija (statistics at the University of Oxford co-advised by Jotun Hein).
D.Phil thesis June 2009.
- Meromit Schuster (computer science).
Ph.D. thesis June 2011.
- Anne Shiu (mathematics).
Ph.D. thesis June 2008.

Graduated students

- Peter Huggins (mathematics), co-advised by Bernd Sturmfels, Ph.D. thesis June 2008.
Polytopes in Computational Biology.
Now a Lane Fellow in Computational Biology at Carnegie Mellon University.
- Radu Mihaescu (mathematics), co-advised by Satish Rao, Ph.D. thesis June 2008.
Distance Methods for Phylogeny Reconstruction.
Winner of the Bernard Friedman Memorial Prize for an outstanding thesis in Applied Mathematics.
- Anat Caspi (bioengineering), Ph.D. thesis June 2007
Comparative Genomics of Repeat Elements.
Now a postdoctoral researcher at the University of Pennsylvania (with David Roos).
- Ariel Schwartz (computer science), Ph.D. December 2006.
Posterior Decoding Methods for Optimization and Accuracy Control of Multiple Alignments.
Now a researcher at Synthetic Genomics.
- Colin Dewey (computer science), Ph.D. June 2006.
Whole-Genome Alignments and Polytopes for Comparative Genomics.
Now an assistant professor of medical informatics and computer science at the University of Wisconsin, Madison.
- Sourav Chatterji (computer science), Ph.D. June 2006.
Computational Analyses of Eukaryotic Gene Evolution.
Now a postdoctoral researcher at UC Davis (with Jonathan Eisen).
- Kevin Chen (computer science), co-advised by Satish Rao, Ph.D. December 2005.
Three Variations on the Theme of Comparative Genomics: Metagenomics, Mitochondrial Gene Rearrangements and MicroRNAs.
Now an assistant professor of genetics and biomathematics at Rutgers University.
- Dan Levy (mathematics), co-advised by Rainer Sachs, Ph.D. December 2005.
Applications of Graph Theory to Chromosome Rearrangements and Phylogenetics,

- Now a postdoctoral researcher at Cold Spring Harbor Laboratories (with Michael Wigler).
- Eric H. Kuo (mathematics and computer science), Ph.D. June 2005.
Combinatorics of Viterbi Sequences.

Former postdocs

- Niko Beerenwinkel (mathematics).
Postdoctoral Research Fellow, 2004–2006.
Now an assistant professor at ETH Zürich, Switzerland.
- Sagi Snir (mathematics and computer science).
Postdoctoral Research Fellow, 2004–2006.
Now an assistant professor at Netanya College, Israel.
- Mathias Drton (mathematics and statistics).
Postdoctoral Research Fellow, 2004–2005.
Now an assistant professor of statistics at the University of Chicago.
- Ruriko Yoshida (mathematics).
Postdoctoral Research Fellow, Summer 2004.
Now an assistant professor of statistics at the University of Kentucky.
- Luis David Garcia-Puente (mathematics).
Postdoctoral Research Fellow, Summer 2004.
Now a visiting assistant professor of mathematics, Texas A & M University.
- Von Bing Yap (mathematics and statistics).
Postdoctoral Research Fellow 2002-2004.
Now an assistant professor of statistics and applied probability, University Singapore.
- Marina Alexandersson (mathematics and statistics).
Postdoctoral Research Fellow 2001-2002.
Now a senior researcher at the Chalmers University of Technology, Sweden.

Other

- Matan Harel (mathematics)
Senior thesis: Split systems and phylogenies, June 2008
- Oscar Westesson (mathematics)
Senior thesis: phylogenetic hidden Markov models, December 2007
- Frances Hammock (mathematics).
Senior thesis: On minors of 4×4 matrices, June 2005 (Ph.D. student UCLA).
- Kushal Chakrabarti (computer science).
Senior thesis: Alignment visualization, June 2004 (Research scientist, amazon.com).
- Nicolas Bray (mathematics).
Senior thesis: Multiple alignment, June 2003 (Ph.D. student, UC Berkeley).
- Eli Rusman (mathematics).
Senior thesis: Genome compression, December 2003 (Engineer, Illumina Inc.).
- Fumei Lam (mathematics).
Senior thesis: Forcing numbers of graphs, June 2000 (Postdoc Brown University).

- Joshua Bao (computer science) and Romanos-Diogenes Malikiosis (mathematics).
Research Summer Institute 1997 (Ph.D. students at Stanford and UCLA).
- Peter Kim (mathematics) and Patience Sethaba (mathematics).
Research Summer Institute 1996 (Ph.D. 2006, Stanford).
- Radoš Radoičić (mathematics).
Summer Program for Undergraduate Research 1997-1998 (Assistant Professor Rutgers).

Doctoral Thesis Committees and Examinations

- Nicolas Bray (Mathematics) – Mathematical biology.
- Robert Bradley (Bioengineering) – RNA folding.
- James Brown (Applied Sciences and Technology) – Algebraic statistics and genomics.
- Anat Caspi (Bioengineering) – Transposable elements.
- Kevin Chen (Computer Science) – Metagenomics.
- Mike Develin (Mathematics) – Topics in discrete geometry.
- Colin Dewey (Computer Science) – Multiple alignment.
- Jeffrey Doker (Mathematics) – Discrete geometry.
- Sourav Chatterji (Computer Science) – Gene finding.
- Isaac Elias (Computer Science) – Phylogenetics
Invited opponent for doctoral thesis defense at KTH, Stockholm.
- Nicholas Eriksson (Mathematics) – Algebraic combinatorics for computational biology.
- Rob Johnson (Computer Science) – Network security.
- Emily Hare (Molecular and Cell Biology) – Gene regulation in *Drosophila*.
- Ben Hsiung (Molecular and Cell Biology) – Regulation of NKG2D ligand expression.
- Peter Huggins (Mathematics) – Polyhedral geometry and computational biology.
- Venky Iyer (Molecular and Cell Biology) – Gene finding in *Drosophila*.
- Hunter Fraser (Molecular and Cell Biology) – Molecular and genomic evolution.
- Anna Kedzierksa (Mathematics) – Comparative genomics
Student at the Universitat Politècnica de Catalunya.
- Eric H. Kuo (Computer Science) – Combinatorics of hidden Markov models.
- Albert Lee (Engineering) – Numerical linear algebra.
- Dan Levy (Mathematics) – Graph theory and phylogenetics.
- Jon McAuliffe (Statistics and Computer Science) – Graphical models.
- Radu Mihaescu (Mathematics) – Phylogenetics.
- Jason Morton (Mathematics) – Algebraic statistics and computational biology.
- Manikandan Narayanan (Computer Science) – Biological network comparison.
- Geoffrey Pike (Computer Science) – Storage optimizations for scientific programs.
- Rahul Satija (Bioinformatics and Statistics) – Statistical alignment.
Student at the University of Oxford.
- Ariel Schwartz (Computer Science) – Multiple alignment.
- Anne Shiu (Mathematics) – Algebraic statistics and computational biology.
- Allan Sly (Statistics) – Glauber dynamics.
- Avinash Varadarajan (Bioengineering) – Statistical alignment.
- Na Xu (Statistics) – Comparative genomics.
- Josephine Yu (Mathematics) – Tropical geometry and combinatorics.
- Xiaoyue Zhao (Statistics) – Modeling of splice sites.

Personal Information

Born May 3, 1973, in Ramat Gan, Israel. Citizen of the United States.
Married to Ingileif Bryndís Hallgrímsdóttir. One daughter.